

**SISTEM TANYA JAWAB EKSTRAKTIF PADA TEKS
BERBAHASA INDONESIA DENGAN *FINE-TUNING* INDOBERT**

*Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-I Pada
Jurusan Teknik Informatika*



Oleh:

Dewa Sheva Dzaky

NIM : 09021182126005

Jurusan Teknik Informatika

FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA

2025

HALAMAN PENGESAHAN

SKRIPSI

SISTEM TANYA JAWAB EKSTRAKTIF PADA TEKS BERBAHASA INDONESIA DENGAN FINE-TUNING INDOBERT

Sebagai salah satu syarat untuk penyelesaian studi di
Program Studi S1 Teknik Informatika

Oleh:

DEWA SHEVA DZAKY

09021182126005

Pembimbing 1 : Dr. Abdiansah, S.Kom., M.Cs.
NIP. 198410012009121005

Mengetahui

Ketua Jurusan Teknik Informatika



Hadipurnawan Satria, Ph.D
19800418202012100

TANDA LULUS UJIAN KOMPREHENSIF

Pada hari Jumat tanggal 09 Mei 2025 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Dewa Sheva Dzaky
Nim : 09021182126005
Judul : Sistem Tanya Jawab Ekstraktif pada Teks Berbahasa Indonesia dengan *Fine-Tuning* IndoBERT

dan dinyatakan LULUS.

1. Ketua Penguji

Al Farissi, M.Cs., Ph.D.
NIP. 198512152014041001


.....

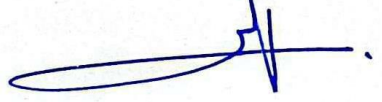
2. Penguji I

Prof. Dr. Ermatita, M.Kom.
NIP. 196709132006042001


.....

3. Pembimbing I

Dr. Abdiansah, S.Kom., M.Cs.
NIP. 198410012009121005


.....

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Dewa Sheva Dzaky

NIM : 09021182126005

Program Studi : Teknik Informatika

Judul Skripsi : Sistem Tanya Jawab pada Teks Berbahasa Indonesia dengan *Fine-Tuning* IndoBERT

Hasil pengecekan Software Turnitin : 4%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan proyek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan dari pihak mana pun.



Indralaya, 21 Mei 2025

Penulis,



Dewa Sheva Dzaky
NIM. 09021182126005

MOTTO DAN PERSEMBAHAN

Motto:

*“Lakukan apa yang kau mau sekarang
Saat hatimu bergerak, jangan kau larang
Hidup ini tak ada artinya, maka
Kau bebas mengarang maknanya seorang”*

- Hindia – Berdansalah, Karir Ini Tak Ada Artinya

Kupersembahkan karya tulis ini kepada:

- Allah SWT
- Orang Tua
- Keluarga Besar
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

ABSTRACT

The abundance of digital information in today's era makes the extraction of relevant information a major challenge, especially in Indonesian, which has unique linguistic characteristics. As an effort to overcome this challenge, this study develops an extractive question-answering system for Indonesian text by fine-tuning the IndoBERT model, which enables the system to extract specific parts of a context paragraph as answers to given questions. The dataset used in this study is the Indonesian-translated version of the Stanford Question Answering Dataset (SQuAD) 2.0, which contains more than 100,000 question-answer pairs derived from Wikipedia articles. The fine-tuning process was carried out in eight scenarios, which are combinations of dataset type (the full dataset including unanswerable questions and a modified dataset with all unanswerable questions removed), learning rate ($2e-5$ and $5e-5$), and batch size (16 and 48). The results of the study show that the model with a learning rate of $5e-5$ and batch size of 16 delivers the best performance. On the dataset with unanswerable questions, the model achieved an exact match score of 60.57% and an f1-score of 70.84%. Meanwhile, on the dataset without unanswerable questions, the model achieved an exact match score of 54.79% and an f1-score of 73.06%.

Keywords: *Extractive Question Answering System, SQuAD, IndoBERT, fine-tuning, exact match, f1-score*

ABSTRAK

Melimpahnya informasi digital di era saat ini menjadikan ekstraksi informasi yang relevan sebagai tantangan besar, terutama dalam bahasa Indonesia yang memiliki karakteristik linguistik yang unik. Sebagai upaya mengatasi tantangan tersebut, penelitian ini mengembangkan sistem tanya jawab ekstraktif pada teks berbahasa Indonesia dengan *fine-tuning* model IndoBERT, yang memungkinkan sistem mengekstraksi bagian tertentu dari sebuah paragraf konteks sebagai jawaban atas pertanyaan yang diajukan. *Dataset* yang digunakan dalam penelitian ini adalah versi terjemahan ke bahasa Indonesia dari *Stanford Question Answering Dataset* (SQuAD) 2.0 yang berisi lebih dari 100.000 pasangan pertanyaan-jawaban yang diambil dari artikel Wikipedia. Proses *fine-tuning* dilakukan dalam delapan skenario, yang merupakan kombinasi dari jenis *dataset* (*dataset* lengkap yang mencakup *unanswerable questions* dan *dataset* yang telah dimodifikasi dengan menghapus semua *unanswerable questions*), *learning rate* ($2e-5$ dan $5e-5$), serta *batch size* (16 dan 48). Hasil penelitian menunjukkan bahwa model dengan *learning rate* $5e-5$ dan *batch size* 16 memberikan performa terbaik. Pada *dataset* dengan *unanswerable questions*, model mencapai nilai *exact match* sebesar 60,57% dan *f1-score* sebesar 70,84%. Sementara itu, pada *dataset* tanpa *unanswerable questions*, model mencapai nilai *exact match* sebesar 54,79% dan *f1-score* sebesar 73,06%.

Kata Kunci: Sistem Tanya Jawab Ekstraktif, SQuAD, IndoBERT, *fine-tuning*, *exact match*, *f1-score*

KATA PENGANTAR

Puji dan syukur kepada Allah SWT yang telah memberikan nikmat sehat, iman dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Skripsi ini dibuat sebagai salah satu syarat untuk menyelesaikan Pendidikan program Strata-1 di Fakultas Ilmu Komputer Universitas Sriwijaya. Dalam proses pembuatan skripsi ini penulis menerima bimbingan, bantuan, semangat, maupun petunjuk dari banyak pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada:

1. Allah SWT atas rahmat dan nikmat-Nya sehingga, penulis dapat menyelesaikan skripsi ini dengan baik.
2. Kedua orang tua dan keluarga yang telah mendoakan, memberi semangat dan motivasi untuk menyelesaikan skripsi ini.
3. Bapak Hadipurnawan Satria, Ph.D. selaku ketua jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Samsuryadi, M.Kom., Ph.D. selaku Dosen dan sekaligus pembimbing akademik yang telah memberikan banyak bantuan dan arahan kepada penulis selama perkuliahan.
5. Bapak Dr. Abdiansah, S.Kom., M.Cs. selaku Dosen Pembimbing skripsi I yang telah memberi bimbingan, arahan serta semangat kepada penulis dalam menyelesaikan skripsi ini.
6. Seluruh dosen program studi serta admin Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

7. Seluruh staf Administrasi dan Pegawai Fakultas Ilmu Komputer yang telah membantu dalam urusan administrasi.
8. Rekan-rekan perjuangan penulis selama perkuliahan, Adit, Anhar, Bima, Epan, Hafiz, Iman, Pandi, Wahyu, Zidane. Terima kasih atas kebersamaan, dukungan, dan saling mengingatkan satu sama lain, terutama dalam proses menyelesaikan skripsi ini. Perjalanan perkuliahan terasa lebih ringan dan bermakna berkat kehadiran kalian.
9. Sahabat-sahabat penulis, Angel, Della, Nisa, Putri, Robby, Shatia, dan Tian yang telah menemani hari-hari penulis selama tinggal di Indralaya. Kehadiran kalian menjadi bagian penting dalam kehidupan sehari-hari, baik dalam suka maupun duka.
10. Pihak-pihak lain yang tidak dapat penulis sebutkan satu-persatu.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak sekali kekurangan karena kurangnya pengalaman dan pengetahuan penulis. Oleh karena itu penulis mengharapkan saran dan kritik yang membangun guna kemajuan penelitian selanjutnya. Semoga tugas akhir ini dapat bermanfaat. Terima Kasih.

Indralaya, 21 Mei 2025

Penulis,



Dewa Sheva Dzaky

DAFTAR ISI

HALAMAN PENGESAHAN SKRIPSI.....	ii
TANDA LULUS UJIAN KOMPREHENSIF	iii
HALAMAN PERNYATAAN	iv
MOTTO DAN PERSEMBAHAN	v
<i>ABSTRACT</i>	vi
ABSTRAK	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xv
BAB I PENDAHULUAN	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang Masalah.....	I-1
1.3 Rumusan Masalah	I-4
1.4 Tujuan Penelitian.....	I-5
1.5 Manfaat Penelitian	I-5
1.6 Batasan Penelitian	I-6
1.7 Sistematika Penulisan	I-7
1.8 Kesimpulan	I-8
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori	II-1
2.2.1 Sistem Tanya Jawab	II-1
2.2.2 Sistem Tanya Jawab Ekstraktif	II-2
2.2.3 <i>Natural Language Processing</i> (NLP)	II-3
2.2.4 <i>Deep Learning</i>	II-3
2.2.5 Transformer	II-4
2.2.6 <i>Bidirectional Encoder Representation from Transformers</i> (BERT) II-10	
2.2.7 IndoBERT	II-14
2.2.8 <i>WordPiece Tokenization</i>	II-15
2.2.9 <i>Fine-Tuning</i>	II-17

2.2.10	<i>Hyperparameter</i>	II-18
2.2.11	Evaluasi Model.....	II-19
2.2.12	<i>Rational Unified Process (RUP)</i>	II-21
2.3	Penelitian Lain yang Relevan.....	II-24
2.4	Kesimpulan	II-26
BAB III METODOLOGI PENELITIAN.....		III-1
3.1	Pendahuluan	III-1
3.2	Pengumpulan Data	III-1
3.2.1	Jenis dan Sumber Data	III-1
3.2.2	Metode Pengumpulan Data	III-3
3.3	Tahapan Penelitian	III-5
3.3.1	Mengumpulkan Data.....	III-5
3.3.2	Menentukan Kerangka Kerja Penelitian	III-6
3.3.3	Menentukan Kriteria Pengujian	III-8
3.3.4	Menentukan Format Data Pengujian.....	III-9
3.3.5	Menentukan Alat Bantu Penelitian.....	III-12
3.3.6	Menentukan Format Data Pengujian.....	III-12
3.3.7	Melakukan Analisis dan Menarik Kesimpulan Penelitian	III-13
3.4	Metode Pengembangan Perangkat Lunak.....	III-13
3.5	Kesimpulan	III-14
BAB IV PENGEMBANGAN PERANGKAT LUNAK		IV-1
4.1	Pendahuluan	IV-1
4.2	Fase Insepsi	IV-1
4.2.1	Pemodelan Bisnis	IV-1
4.2.2	Kebutuhan Sistem	IV-2
4.2.3	Analisis dan Perancangan	IV-2
4.3	Fase Elaborasi	IV-13
4.3.1	Pemodelan Bisnis	IV-13
4.3.2	Kebutuhan Sistem	IV-18
4.3.3	Analisis dan Perancangan	IV-19
4.4	Fase Konstruksi.....	IV-24
4.4.1	Kebutuhan Sistem	IV-24
4.4.2	Implementasi	IV-25
4.5	Fase Transisi.....	IV-29

4.5.1	Pemodelan Bisnis	IV-29
4.5.2	Rencana Pengujian	IV-29
4.5.3	Implementasi	IV-30
4.6	Kesimpulan	IV-32
BAB V HASIL DAN ANALISIS		V-1
5.1	Pendahuluan	V-1
5.2	Hasil Penelitian	V-1
5.2.1	Konfigurasi Pengujian.....	V-1
5.2.2	Hasil Pengujian Skenario 1	V-3
5.2.3	Hasil Pengujian Skenario 2	V-4
5.2.4	Hasil Pengujian Skenario 3	V-6
5.2.5	Hasil Pengujian Skenario 4	V-7
5.2.6	Hasil Pengujian Skenario 5	V-9
5.2.7	Hasil Pengujian Skenario 6	V-10
5.2.8	Hasil Pengujian Skenario 7	V-12
5.2.9	Hasil Pengujian Skenario 8	V-13
5.3	Analisis Hasil Penelitian	V-15
5.4	Kesimpulan	V-21
BAB VI KESIMPULAN DAN SARAN		VI-1
6.1	Pendahuluan	VI-1
6.2	Kesimpulan	VI-1
6.3	Saran.....	VI-2
DAFTAR PUSTAKA		xvii
LAMPIRAN.....		xx

DAFTAR TABEL

Tabel II-1. Contoh Sistem Tanya Jawab Ekstraktif (Islamey et al., 2024).....	II-2
Tabel III-1. Atribut <i>Dataset</i> SQuAD 2.0.....	III-2
Tabel III-2. Contoh Data yang Digunakan.....	III-3
Tabel III-3. Konfigurasi Pengujian	III-10
Tabel III-4. Tabel Contoh Hasil Pengujian.....	III-10
Tabel III-5. Tabel Hasil Pengujian	III-11
Tabel III-6. Alat Bantu Penelitian	III-12
Tabel IV-1. Kebutuhan Fungsional	IV-2
Tabel IV-2. Kebutuhan Non-Fungsional	IV-2
Tabel IV- 3. Contoh Konteks.....	IV-4
Tabel IV-4. Contoh Pertanyaan dengan Atribut <i>is_impossible true</i> dan <i>false</i>	IV-4
Tabel IV-5. Perbedaan Jumlah Data Sebelum dan Sesudah Filter	IV-5
Tabel IV-6. Jumlah Data Setelah Proses Pembagian Data	IV-6
Tabel IV-7. Hasil <i>Encoding</i>	IV-6
Tabel IV-8. Definisi Aktor.....	IV-9
Tabel IV-9. Definisi <i>Use Case</i>	IV-10
Tabel IV-10. Prediksi Jawaban dari Teks	IV-10
Tabel IV-11. Evaluasi Model dengan <i>Dataset</i> Pengujian.....	IV-12
Tabel IV-12. Implementasi Kelas.....	IV-25
Tabel IV-13. Rencana Pengujian <i>Use Case</i> Prediksi.....	IV-29
Tabel IV-14. Rencana Pengujian <i>Use Case</i> Evaluasi	IV-30
Tabel IV-15. Pengujian <i>Use Case</i> Prediksi	IV-30
Tabel IV-16. Pengujian <i>Use Case</i> Evaluasi.....	IV-31
Tabel V-1. Konfigurasi Pengujian	V-1
Tabel V-2. Metrik Evaluasi Skenario 1	V-4
Tabel V-3. Metrik Evaluasi Skenario 2	V-5
Tabel V-4. Metrik Evaluasi Skenario 3	V-7

Tabel V-5. Metrik Evaluasi Skenario 4	V-8
Tabel V-6. Metrik Evaluasi Skenario 5	V-10
Tabel V-7. Metrik Evaluasi Skenario 6	V-11
Tabel V-8. Metrik Evaluasi Skenario 7	V-13
Tabel V-9. Metrik Evaluasi Skenario 8	V-15
Tabel V-10. Hasil Pengujian 8 Skenario	V-15
Tabel V-11. Contoh Data Pengujian	V-19
Tabel V-12. Hasil Pengujian Model Pada Data Sampel.....	V-20

DAFTAR GAMBAR

Gambar II-1. Arsitektur Transformer (Vaswani et al., 2017).....	II-5
Gambar II-2. <i>Scaled Dot-Product Attention</i> (Vaswani et al., 2017).....	II-6
Gambar II-3. <i>Multi-Head Attention</i> (Vaswani et al., 2017).....	II-8
Gambar II-4. <i>WordPiece Tokenization</i> dengan <i>BERT Tokenizer</i> (Karimah & Baita, 2024)	II-16
Gambar II-5. <i>Fine-Tuning</i> pada BERT.....	II-18
Gambar II-6. Dimensi RUP (Prabowo, 2020).....	II-23
Gambar III-1. Tahapan Penelitian	III-5
Gambar III-2. Kerangka Kerja Penelitian	III-6
Gambar IV-1. <i>Use Case Diagram</i>	IV-9
Gambar IV-2. Desain Antarmuka Halaman Beranda	IV-15
Gambar IV-3. Desain Antarmuka Halaman Prediksi	IV-16
Gambar IV-4. Desain Antarmuka Halaman Evaluasi.....	IV-17
Gambar IV-5. <i>Activity Diagram</i> Prediksi Jawaban dari Teks	IV-20
Gambar IV-6. <i>Activity Diagram</i> Evaluasi Model dengan <i>Dataset</i> Pengujian..	IV-21
Gambar IV-7. <i>Sequence Diagram</i> Prediksi Jawaban dari Teks.....	IV-22
Gambar IV-8. <i>Sequence Diagram</i> Evaluasi Model dengan <i>Dataset</i> Pengujian	IV-23
Gambar IV-9. <i>Class Diagram</i>	IV-24
Gambar IV-10. Implementasi Antarmuka Halaman Beranda	IV-26
Gambar IV-11. Implementasi Antarmuka Halaman Prediksi.....	IV-27
Gambar IV-12. Implementasi Antarmuka Halaman Evaluasi.....	IV-28
Gambar V-1. Grafik <i>Loss</i> , <i>Exact Match</i> , dan <i>F1-Score</i> Skenario 1	V-3
Gambar V-2. Grafik <i>Loss</i> , <i>Exact Match</i> , dan <i>F1-Score</i> Skenario 2	V-4
Gambar V-3. Grafik <i>Loss</i> , <i>Exact Match</i> , dan <i>F1-Score</i> Skenario 3	V-6
Gambar V-4. Grafik <i>Loss</i> , <i>Exact Match</i> , dan <i>F1-Score</i> Skenario 4	V-7
Gambar V-5. Grafik <i>Loss</i> , <i>Exact Match</i> , dan <i>F1-Score</i> Skenario 5	V-9
Gambar V-6. Grafik <i>Loss</i> , <i>Exact Match</i> , dan <i>F1-Score</i> Skenario 6	V-10
Gambar V-7. Grafik <i>Loss</i> , <i>Exact Match</i> , dan <i>F1-Score</i> Skenario 7	V-12

Gambar V-8. Grafik <i>Loss</i> , <i>Exact Match</i> , dan <i>F1-Score</i> Skenario 8	V-13
Gambar V-9. Hasil pengujian 8 Skenario.....	V-16

BAB I PENDAHULUAN

1.1 Pendahuluan

Bab ini akan membahas landasan penelitian yang mencakup latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan penelitian, dan sistematika penulisan. Seluruh pembahasan tersebut berperan sebagai kerangka dasar dalam proses penelitian.

1.2 Latar Belakang Masalah

Di era digital saat ini, jumlah informasi digital yang tersedia telah mencapai skala yang belum pernah terjadi sebelumnya. Volume data yang dihasilkan di seluruh dunia diperkirakan mencapai 149 *zettabytes* pada tahun 2024 dan diproyeksikan akan terus meningkat secara eksponensial dalam beberapa tahun ke depan (Taylor, 2024). Data tersebut mencakup berbagai bentuk mulai dari teks, gambar, hingga video yang tersebar di *platform* media sosial, situs berita, dan repositori digital lainnya. Meskipun ketersediaan informasi sangat melimpah, masyarakat umum sering mengalami kesulitan dalam mengekstrak informasi yang relevan dan akurat dari teks. Hal ini disebabkan oleh kompleksitas bahasa, variasi struktur teks, dan kebutuhan akan pemahaman konteks yang mendalam, sehingga menghambat akses informasi yang efisien bagi masyarakat dari berbagai latar belakang.

Sistem tanya jawab ekstraktif (*Extractive Question Answering System*) menjadi salah satu solusi potensial untuk mengatasi tantangan tersebut. Sistem ini memungkinkan ekstraksi bagian tertentu dari sebuah paragraf konteks sebagai

jawaban atas pertanyaan yang diajukan (Pearce et al., 2021), sehingga memudahkan masyarakat umum untuk mendapatkan informasi spesifik tanpa harus membaca seluruh dokumen. Namun, implementasinya pada teks berbahasa Indonesia menghadapi kompleksitas tersendiri karena karakteristik linguistik yang unik, seperti morfologi yang kompleks, variasi dialek, dan struktur sintaksis yang berbeda dari bahasa Inggris yang merupakan bahasa dominan dalam penelitian sistem tanya jawab.

Perkembangan di bidang *Natural Language Processing* (NLP) telah mengalami revolusi signifikan dengan hadirnya model bahasa berbasis arsitektur Transformer seperti BERT (*Bidirectional Encoder Representations from Transformers*). BERT yang diperkenalkan oleh Devlin et al. (2019) telah mengubah pendekatan dalam pemahaman bahasa dengan menggunakan representasi kontekstual dua arah yang memungkinkan model untuk memahami makna kata berdasarkan konteks sebelum dan sesudahnya. Dalam konteks bahasa Indonesia, IndoBERT menjadi salah satu model bahasa berbasis Transformer yang dilatih khusus untuk bahasa Indonesia dan menunjukkan kemampuan yang unggul dalam menangani berbagai tugas pemrosesan bahasa Indonesia.

Penelitian oleh Koto et al. (2020) menunjukkan bahwa IndoBERT secara konsisten melampaui model sebelumnya dalam berbagai tugas NLP untuk bahasa Indonesia, seperti *Named Entity Recognition* (NER) dengan *f1-score* mencapai 90,1%, *Labeled Attachment Score* (LAS) sebesar 82,32% untuk *dependency parsing*, serta akurasi 93,7% dalam *Next Tweet Prediction* (NTP). Meskipun demikian, sumber daya dan model bahasa untuk bahasa Indonesia masih relatif

terbatas dibandingkan dengan bahasa Inggris. Hal ini menyebabkan kesulitan dalam mengekstrak informasi secara akurat dari teks berbahasa Indonesia, terutama dalam konteks sistem tanya jawab ekstraktif yang memerlukan pemahaman mendalam terhadap konteks dan makna teks.

Penelitian terdahulu mengenai sistem tanya jawab berbahasa Indonesia telah menunjukkan perkembangan yang signifikan. Nur Ahmad & Romadhony (2023) menggunakan kombinasi TF-IDF sebagai *retriever* dan IndoBERT sebagai *reader* dalam pengembangan sistem tanya jawab *end-to-end*, yang menghasilkan *exact match* sebesar 53% dan *f1-score* sebesar 70% pada *dataset* terjemahan SQuAD. Sementara itu, penelitian yang dilakukan oleh Masruroh et al. (2023) mengembangkan model tanya jawab berbasis IndoBERT untuk mengevaluasi konten *website* berbahasa Indonesia, dengan hasil *exact match* sebesar 60,21% dan *f1-score* sebesar 71,36%.

Meskipun hasil dari penelitian-penelitian tersebut cukup menjanjikan, masih terdapat ruang untuk peningkatan, terutama dalam menangani kompleksitas bahasa Indonesia dan meningkatkan performa model. *Fine-tuning* merupakan pendekatan yang sangat efektif dalam meningkatkan performa model pada tugas-tugas spesifik. Teknik ini melibatkan penyesuaian parameter model yang telah di-*pre-training* dengan data *domain* tertentu atau tugas spesifik (Liu et al., 2019). Strategi *fine-tuning* yang optimal dapat secara signifikan meningkatkan akurasi dan efisiensi model dalam menangani tugas-tugas NLP yang kompleks, terutama pada sistem tanya jawab ekstraktif yang membutuhkan pemahaman kontekstual yang mendalam.

Keberhasilan IndoBERT dalam berbagai tugas NLP, serta belum adanya *baseline* resmi untuk tugas tanya jawab ekstraktif dalam penelitian sebelumnya, menjadi alasan utama penelitian ini. Dengan melakukan *fine-tuning* IndoBERT untuk sistem tanya jawab ekstraktif pada teks berbahasa Indonesia, diharapkan dapat dikembangkan model yang lebih akurat dan andal dalam mengekstrak jawaban yang relevan dari teks. Model yang dikembangkan ditujukan untuk dapat digunakan oleh masyarakat umum dari berbagai latar belakang dan tingkat pendidikan, tidak terbatas pada kalangan akademisi atau profesional teknologi informasi saja. Dengan demikian, sistem ini dapat menjadi alat bantu yang inklusif dan mudah diakses bagi siapapun yang ingin mendapatkan informasi spesifik dari teks berbahasa Indonesia dengan cepat dan tepat.

Penelitian ini tidak hanya berfokus pada peningkatan performa sistem tanya jawab ekstraktif pada teks berbahasa Indonesia, tetapi juga berkontribusi dalam pengembangan sumber daya pemrosesan bahasa alami untuk bahasa Indonesia yang dapat dimanfaatkan secara luas oleh masyarakat. Hasil yang diperoleh diharapkan dapat menjadi *baseline* baru bagi tugas tanya jawab ekstraktif dalam bahasa Indonesia, mempermudah akses informasi bagi seluruh lapisan masyarakat, dan mendukung penelitian lebih lanjut di bidang ini untuk menghasilkan teknologi yang semakin inklusif dan bermanfaat bagi publik.

1.3 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka didapat rumusan masalah pada penelitian ini yaitu:

1. Bagaimana membuat sistem tanya jawab ekstraktif pada teks berbahasa Indonesia dengan melakukan *fine-tuning* IndoBERT?
2. Bagaimana kinerja sistem tanya jawab ekstraktif pada teks berbahasa Indonesia dengan *fine-tuning* IndoBERT berdasarkan metrik *exact match* dan *f1-score*?

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan, maka terdapat tujuan:

1. Menghasilkan sistem yang dapat menjawab pertanyaan dengan mengekstrak jawaban dari teks berbahasa Indonesia dengan *fine-tuning* IndoBERT.
2. Mengetahui kinerja sistem tanya jawab ekstraktif pada teks berbahasa Indonesia dengan *fine-tuning* IndoBERT berdasarkan metrik *exact match* dan *f1-score*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Bagi peneliti, penelitian ini memberikan pemahaman mendalam tentang penerapan teknik *fine-tuning* pada model bahasa seperti IndoBERT, khususnya dalam sistem tanya jawab ekstraktif pada teks berbahasa Indonesia.
2. Bagi peneliti selanjutnya, hasil penelitian ini diharapkan dapat berkontribusi bagi komunitas NLP di Indonesia serta menjadi referensi dan *baseline* untuk pengembangan sistem tanya jawab ekstraktif yang lebih canggih.

3. Bagi industri, penelitian ini dapat menjadi dasar dalam pengembangan sistem tanya jawab otomatis yang dapat diterapkan pada berbagai bidang, seperti layanan pelanggan, asisten virtual, dan pencarian informasi berbasis teks, sehingga meningkatkan efisiensi dan kualitas layanan.
4. Bagi masyarakat, penelitian ini memberikan manfaat dalam kemudahan akses informasi melalui sistem tanya jawab berbasis teks berbahasa Indonesia, yang dapat digunakan dalam berbagai aplikasi, seperti pendidikan, layanan publik, dan pencarian informasi secara lebih cepat dan akurat.

1.6 Batasan Penelitian

Batasan masalah penelitian ini adalah sebagai berikut:

1. Penelitian ini difokuskan pada tugas tanya jawab ekstraktif dalam bahasa Indonesia.
2. Sistem hanya mendukung pertanyaan yang jawabannya dapat diekstraksi langsung dari konteks, bukan pertanyaan generatif.
3. Dataset yang digunakan terbatas pada hasil terjemahan SQuAD 2.0 berbahasa Indonesia, yang terdiri dari 130.318 data latih dan 11.858 data validasi.
4. Model yang digunakan terbatas pada *indobert-base-uncased* tanpa dilakukan perbandingan dengan model lain.
5. Penyesuaian hyperparameter hanya dilakukan pada *learning rate* dan *batch size*.

6. Evaluasi kinerja model didasarkan pada 2 metrik, yaitu *exact match* dan *f1-score*.

1.7 Sistematika Penulisan

Sistematika penulisan yang digunakan pada penelitian ini mengikuti standar operasional penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya yakni:

BAB I. PENDAHULUAN

Bab ini menjelaskan latar belakang, perumusan masalah, tujuan dari penelitian, manfaat yang diharapkan, batasan-batasan yang ditetapkan, serta sistematika penulisan dalam laporan akhir ini. Bab ini juga memberikan gambaran umum mengenai alur penelitian, mulai dari proses identifikasi masalah hingga penyajian hasil analisis. Penjelasan ini dimaksudkan untuk membantu pembaca memahami struktur dan arah penelitian secara menyeluruh.

BAB II. KAJIAN LITERATUR

Bab ini membahas teori-teori yang menjadi dasar dalam penelitian, termasuk penjelasan mengenai metode IndoBERT, definisi sistem tanya jawab ekstraktif, serta beberapa studi literatur yang relevan dengan topik penelitian ini.

BAB III. METODOLOGI PENELITIAN

Bab ini menguraikan tahapan-tahapan yang dilakukan selama proses penelitian berlangsung, mulai dari metode pengumpulan data hingga metode perancangan perangkat lunak. Setiap tahap dijelaskan secara detail berdasarkan kerangka kerja yang telah disusun.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Bab ini menjelaskan proses pengembangan perangkat lunak, mulai dari tahap analisis kebutuhan pengguna hingga proses pengujian untuk mengevaluasi performa dan hasil dari perangkat lunak yang dikembangkan.

BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini menyajikan hasil yang diperoleh dari penelitian dan analisis yang dilakukan berdasarkan metode serta tahapan yang telah direncanakan. Analisis ini digunakan sebagai dasar dalam menarik kesimpulan dari penelitian.

BAB VI. KESIMPULAN DAN SARAN

Bab ini memuat kesimpulan yang diambil dari pembahasan pada bab-bab sebelumnya serta saran yang diberikan berdasarkan hasil penelitian yang telah dilakukan.

1.8 Kesimpulan

Pada bab ini telah dijelaskan latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan penelitian, serta sistematika penelitian. Pemaparan dalam bab ini bertujuan untuk memberikan gambaran dasar yang menjadi pokok pikiran dalam penelitian sistem tanya jawab ekstraktif pada teks berbahasa Indonesia dengan *fine-tuning* IndoBERT.

DAFTAR PUSTAKA

- Adhim, F. I., Martin, R. F., Budiprayitno, S., & Rahayu, L. P. (2022). Development of Employee Payroll System using Rational Unified Process (RUP) on Odoo Platform. *Applied Technology and Computing Science Journal*, 5(1), 36–43. <https://doi.org/10.33086/atcsj.v5i1.3696>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). *Language Models are Few-Shot Learners* (arXiv:2005.14165). arXiv. <https://doi.org/10.48550/arXiv.2005.14165>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*.
- Dhingra, B., Danish, D., & Rajagopal, D. (2018). Simple and Effective Semi-Supervised Question Answering. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, 582–587. <https://doi.org/10.18653/v1/N18-2092>
- Howard, J., & Ruder, S. (2018). *Universal Language Model Fine-tuning for Text Classification*.
- Islamey, J., Jonathan, V., Nurzaki, M., & Lucky, H. (2024). Comparative Analysis of Encoder-Based Pretrained Models: Investigating the Performance of BERT Variants in Indonesian Question-Answering. *2024 International Conference on Artificial Intelligence, Blockchain, Cloud Computing, and Data Analytics (ICoABCD)*, 309–314. <https://doi.org/10.1109/ICoABCD63526.2024.10704260>
- Karimah, N., & Baita, A. (2024). Multi-Aspect Sentiment Analysis Pada Review Film Menggunakan Metode Bidirectional Encoder Representations From Transformers (BERT). *Komputika : Jurnal Sistem Komputer*, 13(1), 63–72. <https://doi.org/10.34010/komputika.v13i1.11098>
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. *Proceedings of the 28th International Conference on Computational Linguistics*, 757–770. <https://doi.org/10.18653/v1/2020.coling-main.66>
- Lee, K., Salant, S., Kwiatkowski, T., Parikh, A., Das, D., & Berant, J. (2017). *Learning Recurrent Span Representations for Extractive Question*

Answering (arXiv:1611.01436). arXiv.
<https://doi.org/10.48550/arXiv.1611.01436>

Lewis, P., Oğuz, B., Rinott, R., Riedel, S., & Schwenk, H. (2020). *MLQA: Evaluating Cross-lingual Extractive Question Answering* (arXiv:1910.07475). arXiv. <https://doi.org/10.48550/arXiv.1910.07475>

Li, S., Zhao, P., Zhang, H., Sun, X., Wu, H., Jiao, D., Wang, W., Liu, C., Fang, Z., Xue, J., Tao, Y., Cui, B., & Wang, D. (2024). *Surge Phenomenon in Optimal Learning Rate and Batch Size Scaling* (arXiv:2405.14578). arXiv. <https://doi.org/10.48550/arXiv.2405.14578>

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach* (arXiv:1907.11692). arXiv. <https://doi.org/10.48550/arXiv.1907.11692>

Masruroh, S. U., Putri, K. H., Hulliyah, K., Shofi, I. M., & Amrizal, V. (2023). Development Of A Question Answering Model Using The Indobert Algorithm To Evaluate Website Content In Indonesian Language. *2023 11th International Conference on Cyber and IT Service Management (CITSM)*, 1–6. <https://doi.org/10.1109/CITSM60085.2023.10455383>

Nur Ahmad, G., & Romadhony, A. (2023). End-to-End Question Answering System for Indonesian Documents Using TF-IDF and IndoBERT. *2023 10th International Conference on Advanced Informatics: Concept, Theory and Application (ICAICTA)*, 1–6. <https://doi.org/10.1109/ICAICTA59291.2023.10390111>

Pearce, K., Zhan, T., Komanduri, A., & Zhan, J. (2021). *A Comparative Study of Transformer-Based Language Models on Extractive Question Answering* (arXiv:2110.03142). arXiv. <https://doi.org/10.48550/arXiv.2110.03142>

Prabowo, M. (2020). *METODOLOGI PENGEMBANGAN SISTEM INFORMASI*.

Radliya, N. R., Wibawa, J. C., & Rabbi, A. R. (2024). Geographic Information System for Spatial Utilization Licensing of Tabalong Regency Based on Web and Mobile Applications. *International Journal of Research and Applied Technology*.

Rajpurkar, P., Jia, R., & Liang, P. (2018). *Know What You Don't Know: Unanswerable Questions for SQuAD* (arXiv:1806.03822). arXiv. <https://doi.org/10.48550/arXiv.1806.03822>

Rajpurkar, P., Zhang, J., Lopyrev, K., & Liang, P. (2016). *SQuAD: 100,000+ Questions for Machine Comprehension of Text* (arXiv:1606.05250). arXiv. <https://doi.org/10.48550/arXiv.1606.05250>

- Rizqullah, M. R., Purwarianti, A., & Aji, A. F. (2023). QASiNa: Religious Domain Question Answering Using Sirah Nabawiyah. *2023 10th International Conference on Advanced Informatics: Concept, Theory and Application (ICAICTA)*, 1–6. <https://doi.org/10.1109/ICAICTA59291.2023.10390123>
- Sharifani, K., & Amini, M. (2023). *Machine Learning and Deep Learning: A Review of Methods and Applications*. 10(07).
- Taylor, P. (2024). *Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2023, with forecasts from 2024 to 2028*. <https://www.statista.com/statistics/871513/worldwide-data-created/>
- Thaker, N., & Shukla, A. (2020). Python as Multi Paradigm Programming Language. *International Journal of Computer Applications*, 177(31), 38–42. <https://doi.org/10.5120/ijca2020919775>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need* (arXiv:1706.03762). arXiv. <https://doi.org/10.48550/arXiv.1706.03762>
- Yih, W., Chang, M.-W., He, X., & Gao, J. (2015). Semantic Parsing via Staged Query Graph Generation: Question Answering with Knowledge Base. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 1321–1331. <https://doi.org/10.3115/v1/P15-1128>